

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 February 2001 (22.02.2001)

PCT

(10) International Publication Number
WO 01/12798 A2

(51) International Patent Classification⁷: C12N 15/00 (74) Agents: SCHRELL, Andreas et al.; Maybachstrasse 6A, 70469 Stuttgart (DE).

(21) International Application Number: PCT/EP00/08001 (81) Designated States (national): AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW.

(22) International Filing Date: 16 August 2000 (16.08.2000) (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(25) Filing Language: English (26) Publication Language: English

(30) Priority Data: 99116267.8 18 August 1999 (18.08.1999) EP

(71) Applicant (for all designated States except US): SÜD-WESTDEUTSCHE SAATZUCHT [DE/DE]; H. R. Späth, Im Rheinfeld 1-13, D-76437 Rastatt (DE).

(72) Inventors; and (75) Inventors/Applicants (for US only): LÖRZ, Horst [DE/DE]; Ramckeweg 6a, D-22589 Hamburg (DE). DRESSELHAUS, Thomas [DE/DE]; Friedensallee 13, D-22765 Hamburg (DE). SCHREIBER, Daniela [DE/DE]; Trommelstrasse 35, D-22767 Hamburg (DE). HEUER, Sigrid [DE/DE]; Siekreystrasse 3, D-22459 Hamburg (DE).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/12798 A2

(54) Title: MALE STERILE PLANTS

(57) Abstract: The present invention relates to nucleic acid molecules enabling the specific inhibition of a pollen specific regulatory gene. The invention also relates to methods to obtain transgenic plants which show a pollen specific inhibition or overexpression of a gene and thus male sterility. In addition, the invention relates to methods for isolating pollen specific genes.

Male Sterile PlantsDescription

The present invention relates to isolated nucleic acid molecules useful for the production of male sterile plants, in particular male sterile monocotyledonous plants, to vectors containing the nucleic acid molecules, to host cells containing the vectors, to plants, harvest and propagation material containing the host cells, to methods for obtaining them and to methods for isolating such nucleic acid molecules.

The introduction of genes into transgenic plants is considered to have high commercial value. The transfer of heterologous genes or genes of interest into a plant under control of tissue-specific regulatory elements provides a powerful means of conferring selective advantages to plants and to increase their commercial value. The ability to control gene expression is useful for conferring resistance and immunity to certain diseases or to modify the metabolism of a tissue. Plant genetic engineering techniques also prove useful in generating improved plants for plant breeding purposes, such as male sterile plants, in particular plants exhibiting cytoplasmic male sterility (CMS) or nuclear male sterility (NMS) to be used for hybrid breeding.

Hybrid breeding is widely used for corn, rice, and oil seed rape and is becoming ever more important. The most important mechanism to provide plants useful for hybrid breeding is cytoplasmic male sterility. According to this mechanism the anthers of various varieties do not develop functional pollen, which is inherited by cytoplasmic factors and may be restored by dominant restorer genes. In most instances the genes involved are mitochondrial genes, which are maternally inherited. Upon fertilisation with non-restorer varieties, CMS will be maintained in the progeny allowing the indefinite production of seed of the female parent. Another form of producing male sterile plants is to use nuclear male sterility (NMS) mechanisms. However, its use for the production of hybrid seed is severely limited by the inevitable segregation of male fertile progeny when breeding NMS-plants due to its recessive inheritance. Up to now it was not possible to develop a system allowing the production of completely pollen sterile female plants. Thus, there is a great demand for providing NMS-plants for hybrid breeding.

Male sterile plants may be obtained by transforming plants with gene constructs comprising coding sequences for pollen toxic or detrimental proteins under control of pollen specific regulatory elements, for instance ribonucleases under control of an anther specific promoter. Male sterile plants, however, may also be advantageously obtained by specifically inhibiting the formation or function of gene products involved in pollen development and/or function.

In higher plants, such as maize, the male products of meiosis, the microspores, undergo two mitotic cell divisions. The first pollen mitosis generates the large vegetative cell and a smaller generative cell. The generative cell undergoes a second mitotic division forming the two sperm cells. Tricellular and bicellular pollen can be distinguished, depending on whether the second pollen mitosis is performed in the pollen grain or in the pollen tube after germination. The vegetative cell forms the pollen tube upon germination and transports the sperm cells through the transmitting tract of the style to the embryo sac. Pollen tubes enter the embryo sacs through the micropyle and penetrate one of the two synergids whereupon the pollen tube bursts, releasing the sperm cells into the degenerating synergid. During the double fertilisation process one sperm cell fertilises the egg cell and the other sperm cell fertilises the central cell. Maize pollen tubes have to grow up to 30 cm along the style (silk) to reach the embryo sac which is deeply embedded within the ovary (kernel). Although maize pollen tubes reach a growth rate of approximately 1 cm/h, it takes several hours to reach the embryo sac. During this time pollen tubes have to be guided through the transmitting tissue of the style and new cell wall has to be synthesised to sustain the rapidly growing tube tip. Two groups of pollen-expressed genes are distinguished, the so-called early pollen genes, which are most abundant prior to the first pollen mitosis and the late pollen genes, which are thought to have a function during pollen maturation and tube growth.

Several pollen-specific cDNAs have been isolated, representing unknown proteins, cytoskeletal components, poly(A) binding proteins, translation initiation factors, ubiquitin and others (reviewed by Taylor and Hepler, 1997). Several pollen specific nucleic acid sequences are described. It is known for instance that 5' UTRs (untranslated regions) of late pollen expressed genes from tomato are sufficient to alter translation efficiency in pollen (Bate et al., 1996; Curie and McCormick, 1997). The pollen-specific translation initiation factor eIF-4A8 was shown to be phosphorylated during pollen tube growth and might be involved in regulating translation of late pollen genes (op den Camp and Kuhlemeier, 1998).

Furthermore, a pollen MADS box gene DEFH125 of *Antirrhinum* is disclosed by Zachgo et al., 1997. Immunolocalization studies showed increasing DEFH125 expression in maturing stamen and localisation to the cytoplasm of the vegetative cell in mature pollen. Finally, kinases have been described which are expressed at different stages of pollen development (Lee et al., 1996, Estruch et al., 1994; Mu et al., 1994; Li and Gray, 1997; Wilson et al., 1997) and pollen tube growth (Muschietti et al., 1998). However, there still remains the need for an efficient system to engineer plants producing completely sterile pollen.

Thus, it is considered particularly important to develop and provide means and methods that inhibit selectively wild-type pollen function and therefore allow the production of male sterile plants.

Thus, the technical problem underlying the present invention is to provide nucleic acid molecules for use in cloning and expressing pollen specific or pollen abundant genes, in particular for use in monocotyledonous plants which allow the production of male sterile plants.

The present invention solves the technical problem underlying the present invention by providing purified nucleic acid molecules for use in cloning and expressing a pollen specific or pollen abundant gene in a plant which are selected from the group consisting of

(a) the nucleic acid sequence defined in any one of SEQ ID No. 1 or 12, or part or a complementary strand thereof,

(b) a nucleic acid sequence encoding a protein or peptide with the amino acid sequence defined in SEQ ID No. 2, or part or a complementary strand thereof,

(c) a nucleic acid sequence which hybridises to the nucleic acid sequence defined a) or b), or part or a complementary strand thereof and

(d) a nucleic acid sequence which is degenerate as a result of the genetic code to the nucleic acid sequence defined in a), b), c), or part or a complementary strand thereof,

(e) alleles or derivatives of the nucleic acid sequence defined in (a), (b), (c), (d), or part or a complementary strand thereof.

The present invention also solves this problem by providing purified nucleic acid molecules for use in cloning and expressing a pollen specific or pollen abundant gene in a plant which are selected from the group consisting of

(a) the nucleic acid sequence defined in any one of SEQ ID No. 3 to 11, or a part or complementary strand thereof,

(b) a nucleic acid sequence which hybridises to the nucleic acid sequence defined in (a), or a part or complementary strand thereof and

(c) alleles or derivatives of the nucleic acid sequence defined in (a) or (b), or part or a complementary strand thereof.

The nucleic acid sequence set out in SEQ ID No. 1 represents a nucleic acid sequence, namely a cDNA sequence encoding a pollen specifically expressed protein, called the ZmMADS2 protein, which is probably essential for pollen tube growth. This sequence will be termed in the following the coding sequence of the present invention or the ZmMADS2 coding sequence.

The amino acid sequence set out in SEQ ID No. 2 represents the amino acid sequence of the protein ZmMADS2.

The sequences set out in SEQ ID No. 3 to 9 comprise isolated 5' regulatory elements and are considered to be promoters or parts thereof, i.e. promoter fragments. These sequences are capable of modulating, initiating and/or contributing to the pollen specific or pollen abundant transcription of nucleic acid sequences operably linked to them. In a preferred embodiment of the present invention, these sequences may additionally contain at their 3' terminus the nucleotide sequence from position +1 to 310 as depicted in SEQ ID No. 10. This sequence is a 5' untranslated region, which may serve as linker between a promoter of the present invention and a coding sequence linked thereto.

The sequences set out in SEQ ID No. 11 comprise an isolated 3' regulatory element and is considered to be a 3' transcription regulatory element, or a part thereof, for instance a termination and/or polyadenylation signal. This sequence is capable of modulating, in particular contributing to, or terminating the transcription of nucleic acid sequences operably linked to them.

The sequence set out in SEQ ID No. 12 is the sequence of a genomic clone comprising all of the above sequences.

The present invention also relates to nucleic acid sequences which hybridise, in particular under stringent conditions, to the sequences set out in any one of SEQ ID No. 1 and 3 to 12. In particular, these sequences have a degree of identity of 70% to the sequence of SEQ ID Nos. 1 and 3 to 12.

In the context of the present invention, nucleic acid sequences which hybridise to any one of the specifically disclosed sequences of SEQ. Id. Nos. 1 and 3 to 12 are sequences which have a degree of 60% to 70% sequence identity to the specifically disclosed sequence on nucleotide level. In an even more preferred embodiment of the present invention, sequences which are encompassed by the present invention are sequences which have a degree of identity of more than 70%, and even more preferred, more than 80%, 90%, 95% and particularly 99% to the specifically disclosed sequences on nucleotide level.

Thus, the present invention relates to nucleic acid sequences, in particular DNA sequences which hybridise under the hybridisation conditions as described in Sambrook et al., (1989) in particular under the following conditions to the sequences specifically disclosed:

Hybridisation buffer: 1 M NaCl; 1% SDS; 10% dextran sulphate; 100 µg/ml ssDNA
Hybridisation temperature: 65° C
First wash: 2 x SSC; 0.5% SDS at room temperature
Second wash: 0.2 x SSC; 0.5% SDS at 65°C.

More preferably, the hybridisation conditions are chosen as identified above, except that a hybridisation temperature and second wash temperature of 68° C, and even more preferred, a hybridisation temperature and second wash temperature of 70° C is applied.

Thus, the present invention also comprises nucleic acid sequences which are functionally equivalent to the sequences of any one of SEQ ID No. 1 and 3 to 12, in particular sequences which have at least homology to the sequence of SEQ ID No. 1 and 3 to 12. The invention also relates to alleles and derivatives of the sequences mentioned above which are defined as sequences being essentially similar to the above sequences but comprising, for instance, nucleotide exchanges, substitutions (also by unusual nucleotides), rearrangements, mutations, deletions, insertions, additions or nucleotide modifications and are functionally equivalent to the sequences set out in SEQ ID No. 1 and 3 to 12.

The nucleic acid molecules of the present invention are, in a preferred embodiment, derived from maize (*Zea mays*), most preferably from maize pollen.

According to the present invention it was found that the nucleic acid sequence isolated are specifically expressed in mature pollen and obviously play an important role in pollen development, in particular pollen tube growth, pollen development and/or fertilisation.

Thus, the nucleic acid molecules of the present invention are useful for cloning tissue specific, in particular pollen specific nucleic acid sequences, in particular regulatory elements, coding sequences and/or complete genes, in plants, in particular in monocotyledonous plants. Thus, the present invention provides the means for the isolation of pollen specific coding sequences and/or transcription

regulatory elements that direct or contribute to pollen-preferred gene expression in plants, in particular in monocotyledonous plants, such as maize. The present invention also provides the means of isolating pollen specifically expressed genes and their transcripts.

The nucleic acid molecules of the present invention are also useful for expressing a pollen specific protein, namely the ZmMADS2 coding sequence, in plants, in particular in the pollen of plants and especially in the pollen of monocotyledonous plants such as maize or of dicotyledonous plants such as sugar beets (*Beta vulgaris*). Thus, the present invention provides the means to allow the expression of a particular pollen specific or pollen abundant gene in pollen thereby enabling the modification of pollen development, pollen tube growth or fertilisation processes. The coding sequence of the present invention may be overexpressed in transformed plants due to expression under control of a strong constitutive or regulated promoter. It is also possible to modify the coding sequence of the present invention so as to allow the production of a modified pollen specific protein, which in turn modifies in a desired manner pollen development and/or function. Most importantly, the present invention provides the means to specifically inhibit the formation of a protein essential for pollen function, namely the ZmMADS2 protein, by transforming plants with antisense constructs comprising all or part of the ZmMADS2 transcribed region, in particular the coding sequence or part thereof in an antisense orientation under the control of its wild-

type or appropriate other regulatory elements so as to effectively bind to wild-type ZmMADS2 mRNA and inhibit its translation. Another possibility of eliminating the expression of an endogenous ZmMADS2 gene is to use cosuppression technology, thus to insert a transgenic copy of the endogenous ZmMADS2 gene into the plant to silence the expression, for instance by overexpression of the transgenic copy coupled in sense orientation to a strong promoter. Such a construct leads upon expression to the abolishment of the wild-type ZmMADS2 function thereby producing nuclear male sterile plants. Of course, the present invention also relates to processes to restore the elimination effect, for instance by using switchable or inducible promoters to restore the effects.

The present invention also provides regulatory elements, such as promoters, 3' termination signals or elements present in 5' and/or 3' untranslated regions (UTRs) providing for pollen specific expression of any gene of interest, including the ZmMADS2 coding sequence of the present invention.

In a particularly preferred embodiment of the present invention, the promoter of the present invention is expressed in a spatially and temporarily specific manner, preferably in mature pollen after dehiscence, but before or during dehydration of the pollen takes place. In a further preferred embodiment the promoter of the present invention is able, due to specific sequence elements present in its sequence, not only to direct expression in preferably mature, ungerminated pollen tissue but addi-

tionally also in root tips. Of course, by deleting or inactivating the root specific sequence element in the promoter of the present invention an exclusive pollen specific expression without any expression in other tissues of a plant may be obtained. Accordingly, the proteins encoded by the gene of interest can be accumulated in pollen. For instance, the promoter of the present invention is particularly useful in driving the pollen specific transcription of heterologous structural genes that confer male sterility to transgenic plants by expressing exclusively in pollen mRNA or proteins inhibiting or adversely influencing pollen development and/or function. Thus, the present invention relates to improved means and methods for plant breeding, in particular the production of hybrid seed, for instance by creating plants exhibiting nuclear male sterility, in particular in grain, cereals and corn. In a particularly preferred embodiment, the present invention relates to a DNA construct with a promoter and/or 3' regulatory element of the present invention operably linked to a coding sequence for a toxic protein, for instance a barnase or a coding sequence of the present invention as set out in SEQ ID No. 1.

In the context of the present invention, a number of terms shall be utilised as follows.

The term "promoter" refers to a sequence of DNA, usually upstream (5') to the coding sequence of a structural gene, which controls the expression of the coding region by providing the recognition for RNA polymerase and/or other factors required for

transcription to start at the correct site. Promoter sequences are necessary, but not always sufficient to drive the expression of the gene.

A "3' regulatory element (or "3' end") refers to that portion of a gene comprising a DNA segment, excluding the 5' sequence which drives the initiation of transcription and the structural portion of the gene, that determines the correct termination site and contains a polyadenylation signal and any other regulatory signals capable of effecting messenger RNA (mRNA) processing or gene expression. The polyadenylation signal is usually characterised by effecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. Polyadenylation signals are commonly recognised by the presence of homology to the canonical form 5'-AATAAA-3', although variations are not uncommon.

"Nucleic acid" refers to a large molecule which can be single or double stranded, composed of monomers (nucleotides) containing a sugar, phosphate and either a purine or pyrimidine. The nucleic acid may be cDNA, genomic DNA, or RNA, for instance mRNA.

The term "nucleic acid sequence" refers to a natural or synthetic polymer of DNA or RNA which may be single or double stranded, alternatively containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers.

The term "gene" refers to a DNA sequence that codes for a specific protein and regulatory elements controlling the expression of this DNA sequence.

The term "regulatory element" refers to a sequence located upstream (5'), within and/or downstream (3') to a coding sequence whose transcription and expression is controlled by the regulatory element, potentially in conjunction with the protein biosynthetic apparatus of the cell. "Regulation" or "regulate" refer to the modulation of the gene expression induced by DNA sequence elements located primarily, but not exclusively upstream (5') from the transcription start of the gene of interest. Regulation may result in an all or none response to a stimulation, or it may result in variations in the level of gene expression.

The term "coding sequence" refers to that portion of a gene encoding a protein, polypeptide, or a portion thereof, and excluding the regulatory sequences which drive the initiation or termination of transcription.

The coding sequence or the regulatory element may be one normally found in the cell, in which case it is called "autologous", or it may be one not normally found in a cellular location, in which case it is termed "heterologous".

A heterologous gene may also be composed of autologous elements arranged in an order and/or orientation not normally found in the cell in which it is transferred. A heterologous gene may be derived in whole or in part from any source known to the art, including a bacterial or viral genome or episome, eukaryotic nuclear or plasmid DNA, cDNA or chemically synthesised DNA. The structural gene may

constitute an uninterrupted coding region or it may include one or more introns bounded by appropriate splice junctions. The structural gene may be a composite of segments derived from different sources, naturally occurring or synthetic.

The term "vector" refers to a recombinant DNA construct which may be a plasmid, virus, or autonomously replicating sequence, phage or nucleotide sequence, linear or circular, of a single or double stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and a DNA sequence for a selected gene product in sense or antisense orientation along with appropriate 3' untranslated sequence into a cell, in particular a plant cell.

As used herein, "plant" refers to photosynthetic organisms, such as whole plants including algae, mosses, ferns and plant-derived tissues. "Plant derived tissues" refers to differentiated and undifferentiated tissues of a plant, including pollen, pollen tubes, pollen grains, roots, shoots, shoot meristems, coleoptilar nodes, tassels, leaves, cotyledondous petals, ovules, tubers, seeds, kernels and various forms of cells in culture such as intact cells, protoplasts, embryos and callus tissue. Plant-derived tissues may be in planta, or in organ, tissue or cell culture. A "monocotyledonous plant" refers to a plant whose seeds have only one cotyledon, or organ of the embryo that stores and absorbs food. A

"dicotyledonous plant" refers to a plant whose seeds have two cotyledons.

"Transformation" and "transferring" refers to methods to transfer DNA into cells including, but not limited to, biolistic approaches such as particle bombardment, microinjection, permeabilising the cell membrane with various physical (e.g., electroporation) or chemical (e.g., polyethylene glycol, PEG) treatments; the fusion of protoplasts or Agrobacterium tumefaciens or rhizogenes mediated transformation. For the injection and electroporation of DNA in plant cells there are no specific requirements for the plasmids used. Plasmids such as pUC derivatives can be used. If whole plants are to be regenerated from such transformed cells, there should be a selectable marker. Depending upon the method for the introduction of desired genes into the plant cell, further DNA sequences may be necessary; if, for example, the Ti or Ri plasmid is used for the transformation of the plant cell, at least the right border, often, however, the right and left border of the Ti and Ri plasmid T-DNA have to be linked as flanking region to the genes to be introduced.

If Agrobacteria are used for the transformation, the DNA to be introduced has to be cloned into specific plasmids, either into an intermediary vector or into a binary vector. The intermediary vectors can be integrated into the Ti or Ri plasmid of the Agrobacteria due to sequences that are homologous to sequences in the T-DNA by homologous recombination.

tion. The Ti or Ri plasmid furthermore contains the vir region necessary for the transfer of the T-DNA into the plant cell. Intermediary vectors cannot replicate in Agrobacteria. By means of a helper plasmid the intermediary vector can be transferred by means of a conjugation to *Agrobacterium tumefaciens*. Binary vectors can replicate both in *E.coli* and in Agrobacteria and they contain a selection marker gene and a linker or polylinker framed by the right and left T-DNA border region. They can be transformed directly into the Agrobacteria (Holsters et al., 1978). The Agrobacterium serving as a host cell should contain a plasmid carrying a vir region. The Agrobacterium transformed is used for the transformation of plant cells. The use of T-DNA for the transformation of plant cells has been extensively examined and described in EP-A 120 516; Hoekema, (1985); An et al., (1985).

For the transfer of the DNA into the plant cell plant explants can be co-cultivated with *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. From the infected plant material (e.g., pieces of leaf, stem segments, roots, but also protoplasts or plant cells cultivated by suspension) whole plants can be regenerated in a suitable medium, which may contain antibiotics or biozides for the selection of transformed cells.

Alternative systems for the transformation of monocotyledonous plants are the transformation by means of electrically or chemically induced introduction of DNA into protoplasts, the electroporation of partially permeabilised cells, the macroinjection

of DNA into flowers, the microinjection of DNA into micro-spores and pro-embryos, the introduction of DNA into germinating pollen and the introduction of DNA into embryos by swelling (Potrykus, *Physiol. Plant* (1990), 269-273).

While the transformation of dicotyledonous plants via Ti plasmid vector systems with the help of *Agrobacterium tumefaciens* is well-established, more recent research work indicates that also monocotyledonous plants are accessible for transformation by means of vectors based on *Agrobacterium* (Chan et al., (1993); Hiei et al., (1994); Bytebier et al., (1987); Raineri et al., (1990), Gould et al., (1991); Mooney et al., (1991); Lit et al., (1992)).

In fact, several of the above-mentioned transformation systems could be established for various cereals: the electroporation of tissues, the transformation of protoplasts and the DNA transfer by particle bombardment in regenerative tissue and cells (Jähne et al., (1995)). The transformation of maize has been described in Brettschneider et al., (1997) and Ishida et al., (1996). The transformation of wheat has been frequently described in the literature (Maheshwari et al., (1995)).

The term "host cell" refers to a cell which has been genetically modified by transfer of a heterologous or autologous nucleic acid sequence or its descendants still containing this sequence. The host cell may be transiently or stably transformed and is preferably able to express the transformed nucleic acid molecule. These cells are also

termed "transgenic cells". In the case of an autologous nucleic acid sequence being transferred, the sequence will be present in the host cell in a higher copy number than naturally occurring.

The term "operably linked" refers to the chemical fusion of two or more fragments of DNA in a proper orientation such that the fusion preserves or creates a proper reading frame, or makes possible the proper regulation of expression of the DNA sequences when transformed into plant tissue.

The term "expression" as used herein is intended to describe the transcription and/or coding of the sequence for the gene product. In the expression, a DNA chain coding for the sequence of gene product is first transcribed to a complementary RNA, which is often an mRNA, and then the thus transcribed mRNA is translated into the above mentioned gene product if the gene product is a protein. However, expression also includes the transcription of DNA inserted in antisense orientation to its regulatory elements. Expression, which is constitutive and possibly further enhanced by an externally controlled promoter fragment thereby producing multiple copies of mRNA and large quantities of the selected gene product, may also include overproduction of a gene product.

A "tissue specific promoter" refers to a sequence of DNA that provides recognition signals for RNA polymerase and/or other factors required for transcription to begin, and/or for controlling expression of the coding sequence precisely within cer-

tain tissues or within certain cells of that tissue. Expression in a tissue specific manner may be only in individual tissues, or cells within tissues, or in combinations of tissues. The present invention relates in particular to pollen and/or root specific expression, i.e. examples may include tissue specific expression in pollen only and no other tissues within the plant, or may be in pollen and roots, and no other tissues of the plant. An expression in pollen according to which the expression takes place mainly, but not exclusively, in the pollen is also termed "tissue abundant".

The term "pollen specific nucleic acid sequence" refers to nucleic acid sequences, i.e. genes, coding sequences and/or regulatory elements which are exclusively or mainly active in pollen of plants, in particular those which direct or contribute to a pollen abundant or pollen selective expression of a protein. The term "pollen abundant nucleic acid sequence" refers to nucleic acid sequences, i.e. genes, coding sequences and/or regulatory elements which are mainly active in pollen of plants, in particular those which direct or contribute to a pollen abundant expression of a protein.

In a further preferred embodiment the invention relates to nucleic acid molecules specifically hybridising to transcripts of the nucleic acid molecules. These nucleic acid molecules are preferably oligonucleotides having a length of at least 10, in particular of at least 15 and particularly preferred of at least 50 nucleotides. The nucleic acid molecules and oligonucleotides of the present in-

vention may be used for instance as primers for a PCR reaction or be used as components of antisense constructs or of DNA molecules encoding suitable ribozymes.

The present invention also relates to vectors comprising the above-identified nucleic acid molecules in particular comprising chimeric DNA constructs or non-chimeric DNA constructs such as the wild-type ZmMADS2 gene, or derivatives thereof or parts thereof. The term DNA construct refers to a combination of at least one regulatory element and a coding sequence.

Thus, the present invention relates to recombinant nucleic acid molecules useful in the preparation of male sterile plant cells and plants as defined above by genetic engineering. In particular, the invention concerns chimeric DNA constructs comprising a coding DNA sequence coding for a wild-type ZmMADS2 protein operably linked to a promoter wherein said promoter is different to the promoter linked to said ZmMADS2 coding sequence in the wild-type gene i.e. either is a mutated wild-type promoter or a promoter from another gene and/or species. In a further preferred embodiment, the invention concerns chimeric DNA constructs comprising a modified coding DNA sequence coding for a mutated ZmMADS2 protein, wherein the DNA-sequence is operably linked to a promoter which may be different from the promoter linked to said ZmMADS2 coding sequence in the wild-type gene or the promoter is the wild-type ZmMADS2 promoter.

Of course, the present invention also relates to chimeric antisense constructs comprising a DNA sequence present in or derived from the transcribed region of the ZmMADS2 sequence of the present invention depicted in SEQ ID No 1, or a part thereof, which is linked to a promoter wherein said promoter is different to the promoter linked to said ZmMADS2 coding sequences in the wild-type gene or is the wild-type promoter and wherein the orientation of the coding sequence to the promoter is vice versa to the wild-type orientation. In one embodiment of the present invention, the DNA sequence of the present invention used to specifically inhibit via antisense constructs the translation of ZmMADS2 expression from the wild-type gene is at least partially not derived from the ZmMADS2 coding sequence, but rather contains sequences from untranslated regions of the ZmMADS2 transcribed region. Of course the invention also relates to DNA constructs comprising a DNA sequence coding for the non-chimeric wild-type ZmMADS2 protein operably linked to the wild-type promoter. These constructs may be used to transform plant cells and plants for which the DNA construct is autologous, i.e. is the source or natural environment for the DNA construct, or for which the DNA construct is heterologous, i.e., is from another species. Plant cells and plants obtained by using the above listed DNA constructs may be characterised by ZmMADS2 antisense expression, multiple copies of the above DNA constructs in their genome, that means are characterised by an increased copy number of the ZmMADS2 gene in the genome and/or a different location in the genome with respect to the wild-type gene

and/or the presence of a foreign gene in their genome.

In the context of the present invention a chimeric DNA construct is thus a DNA sequence composed of different DNA fragments not naturally occurring in this combination. The DNA fragments combined in the chimeric DNA construct may originate from the same species or from different species. For example a DNA fragment coding for an ZmMADS2 protein may be operably linked to a DNA fragment representing a promoter from another gene of the same species that provides for an increased expression of the ZmMADS2 coding sequence. Preferably however, a DNA fragment coding for an ZmMADS2 protein is operably linked to a DNA fragment containing a promoter from another species for instance from another plant species, from a fungus, yeast or from a plant virus or a synthetically produced promoter. A synthetically produced promoter is either a promoter synthesised chemically from nucleotides de novo or a hybrid-promoter spliced together by combining two or more nucleotide sequences from synthetic or natural promoters which are not present in the combined form in any organism. The promoter has to be functional in the plant cell to be transformed with the chimeric DNA construct.

The promoter used in the present invention may be derived from the same or from a different species and may provide for constitutive or regulated expression, in particular positively regulated by internal or external factors. External factors for the regulation of promoters are for example light,

draught, heat, chemicals such as inorganic salts, heavy metals or organic compounds such as organic acids, derivatives of these acids, in particular its salts and sugars.

Examples of promoters to be used in the context of the present invention are the cauliflower mosaic virus (CaMV) 19S or 35S promoters, nopaline synthase promoters, pathogenesis-related (PR) protein promoters, the ubiquitin promoter from maize for a constitutive expression, the HMG promoters from wheat, promoters from Zein genes from maize, small subunit of ribulose bisphosphonate carboxylase (ssuRUBISCO) promoters, the 35S transcript promoter from the figworm mosaic virus (FMV 35S), the octopine synthase promoter or the actin promoter from rice etc. It is preferred that the particular promoter selected should be capable of causing sufficient expression to result in the production of an effective amount of antisense mRNA or modified or wild-type ZmMADS2 protein to produce male sterility. Of course, for selective expression of the ZmMADS2 protein tissue specific promoters may be used. However, in the most preferred embodiment of the present invention, i.e. the ZmMADS2 antisense constructs, the promoter may be a constitutive strong promoter, since the pollen specificity of the antisense action is confined to the pollen due to pollen specific expression of the target, i.e. the wild-type ZmMADS2 expression.

The DNA construct of the invention may contain multiple copies of a promoter and/or multiple copies of the DNA coding sequences. In addition, the con-

struct may include coding sequences for markers and coding sequences for other peptides such as signal or transit peptides or resistance genes for instance against virus infections or antibiotics.

Useful markers are peptides providing antibiotic or drug resistance for example resistance to phosphotrycine, hygromycin, kanamycin, G418, gentamycin, lincomycin, methotrexate or glyphosate. These markers, such as the herbicide resistance gene pat encoding a phosphotrycine acetyl transferase, can be used to select cells transformed with the chimeric DNA constructs of the invention from untransformed cells. Of course other markers are markers coding peptidic enzymes which can be easily detected by a visible reaction for example a colour reaction for example luciferase, β -1,3-glucuronidase or β -galactosidase.

Signal or transit peptides provide the ZmMADS2 protein formed on expression of the DNA constructs of the present invention with the ability to be transported to the desired site of action. Examples for transit peptides of the present invention are chloroplast transit peptides, mitochondria transit peptides or nuclear localisation signals.

In chimeric DNA constructs containing coding sequences for transit peptides these sequences are usually derived from a plant, for instance from corn, potato, Arabidopsis or tobacco. Preferably, transit peptides and ZmMADS2 coding sequences are derived from the same plant, for instance corn. In particular such a chimeric DNA construct comprises a DNA sequence coding for a wild-type ZmMADS2 pro-

tein and a DNA sequence coding for a transit peptide operably linked to a promoter wherein said promoter is different to the promoter linked to said coding sequences in wild-type gene, but functional in plant cells. In particular, said promoter provides for higher transcription efficiency than the wild-type promoter.

The mRNA produced by a DNA construct of the present invention may advantageously also contain a 5' non-translated leader sequence. This sequence may be derived from the promoter selected to express the gene and can be specifically modified so as to increase translation and stability of the mRNA. The 5' non-translated regions can also be obtained from viral RNAs from suitable eucaryotic genes or a synthetic gene sequence.

Preferably, the coding sequence of the present invention is not only operably linked to 5' regulatory elements, such as promoters, but is additionally linked to other regulatory elements such as enhancers and/or 3' regulatory elements. For instance, the vectors of the present invention may contain functional terminator sequences such as the terminator of the octopine synthase gene from *Agrobacterium tumefaciens*. Further 3' non-translated regions to be used in a chimeric construct of the present invention to cause the addition of polyadenylate nucleotides to the 3' end of the transcribed RNA are the polyadenylation signals of the *Agrobacterium tumefaciens* nopaline synthase gene (NOS) or from plant genes like the soy bean storage protein gene and the small subunit of the ribulose-

1,5-bisphosphonate carboxylase (ssuRUBISCO) gene. Of course, also the regulating elements of the present invention deriving from the wild-type ZmMADS2 gene may be used.

The vectors of the present invention may also possess functional units effecting the stabilisation of the vector in the host organism, such as bacterial replication origins. Furthermore, the chimeric DNA constructs of the present invention may also encompass introns or part of introns inserted within or outside the coding sequence for the ZmMADS2 protein.

In a particularly preferred embodiment of the present invention the vector furthermore contains T-DNA, in particular the left, the right or both T-DNA borders derived from *Agrobacterium tumefaciens*. Of course, sequences derived from *Agrobacterium rhizogenes* may also be used. The use of T-DNA sequences in the vector of the present invention enables the *Agrobacterium* mediated transformation of cells. In a preferred embodiment of the present invention the nucleic acid sequence of the present invention, optionally operably linked to regulatory elements, is inserted within the T-DNA or adjacent to it.

Furthermore, the present invention relates to a wild-type or modified ZmMADS2 protein coded by a nucleic acid sequence of the present invention. The ZmMADS2 protein exhibits in a particularly preferred embodiment features of a MADS-box protein and in particular transcriptional regulative activity during pollen development, in particular pollen

tube growth, and function. In particular, the present invention relates to a ZmMADS2 protein produced from a plant cell or plant of the present invention or from the propagation material or harvest products of plants or plant cells of the present invention. The invention also relates to antibodies, in particular mono- or polyclonal antibodies as well as peptide antibodies raised against the protein with the activity of an ZmMADS2 protein which may be useful for cloning and detection assays.

Thus, the present invention also relates to a method for the production of a protein with the activity of an ZmMADS2 protein, wherein a cell of the present invention, in particular a plant cell or plant callus is cultivated under conditions allowing the synthesis of the protein and the protein is isolated from cultivated cells and/or the culture medium.

In a particularly preferred embodiment of the present invention the 5' and/or 3' regulatory elements of the present invention contained in the vector are operably linked to a gene of interest which in this context may also be only its coding sequence, which may be a heterologous or autologous gene or coding sequence. Such a gene of interest may be a gene, in particular its coding sequence, conferring, for instance, disease resistance; drought resistance; insect resistance; herbicide resistance; immunity; an improved intake of nutrients, minerals or water from the soil; or a modified metabolism in the plant, particularly its pollen and/or roots.

Such a modified metabolism may relate to a preferred accumulation of useful or toxic substances in pollen and/or roots, for instance sugars or, vice versa, in the depletion of substances undesirable in pollen and/or roots, for instance certain amino acids. Thus, in the context of the present invention, a gene of interest may confer resistance to infection by a virus, such as a gene encoding the capsid protein of the BWYV or the BNYVV virus, a gene conferring resistance to herbicides such as Basta®, or to an insecticide, a gene conferring resistance to the corn rootworm, a gene encoding the toxic crystal protein of *Bacillus thuringiensis* or a gene whose expression confers male sterility. A gene of interest includes also a coding sequence cloned in antisense orientation to the regulatory sequences directing its expression. Such an antisense-construct may be used specifically to repress the activity of undesirable genes in plant cells, in particular in pollen and/or roots, for instance to produce male sterile plants exhibiting a modified, for instance abortive pollen development, pollen morphology or metabolism and/or pollen function. The gene of interest may also comprise signal sequences, in particular ER targeting sequences, directing the encoded protein in the ER and eventually for instance in the cell wall, vascular tissue, nucleus and /or the vacuole.

Thus, the nucleic acid sequences of the present invention are useful since they enable the pollen and/or root specific expression of a ZmMADS2 coding sequence in antisense orientation and further genes of interest in plants, in particular monocotyledon-

dous plants. Accordingly, plants are enabled to produce useful, in particular toxic, products in their pollen or the plants may be engineered by modifying the pollen structure, function and/or development.

The present invention also relates to a method of genetically modifying a cell by transforming it with a nucleic acid molecule of the present invention or vector according to the above, whereby the *ZmMADS2* coding sequence or a further gene of interest operably linked to at least one regulatory element either according to the present invention or as conventionally used is expressible in the cell. In particular, the cell being transformed by the method of the present invention is a plant, bacterial or yeast cell. In a particularly preferred embodiment of the present invention, the above method further comprises the regeneration of the transformed cell to a differentiated and, in a preferred embodiment, fertile or non-fertile plant.

The present invention also relates to host cells transformed with the nucleic acid molecule or the vector of the present invention, in particular plant, yeast or bacterial cells, in particular monocotyledonous or dicotyledonous plant cells. The present invention also relates to cell cultures, tissue, calluses, etc. comprising a cell according to the above, i.e. a transgenic cell and its descendants harbouring and preferably experiencing the nucleic acid molecule or vector of the present invention.

Thus, the present invention relates to transgenic plant cells which were transformed with one or several nucleic acid molecules of the present invention as well as to transgenic plants cells originating from such transformed cells. Such plant cells can be distinguished from naturally occurring plant cells by the observation that they contain at least one nucleic acid molecule according to the present invention which does not naturally occur in these cells, or by the fact that such a molecule is integrated into the genome of the cell at a location where it does not naturally occur, that is, in another genomic region, or by the observation that the copy number of the nucleic acid molecules is different from the copy number in naturally occurring plants, in particular a higher copy number. Preferably, the nucleic acid sequences of the present invention is operably linked in antisense orientation.

Thus, the present invention also relates to transgenic cells, also called host cells, transformed with the nucleic acid molecule or vector of the present invention, in particular plant, yeast or bacterial cells, in particular monocotyledonous or dicotyledonous plant cells. The present invention also relates to cell cultures, tissue, calluses, propagation and harvest material, pollen, seeds, seedlings, embryos, etc. comprising a cell according to the above, that is, a transgenic cell being stably or transiently transformed and being capable of expressing a male sterility causing coding sequence. The transgenic plants of the present invention can be regenerated to whole plants accord-

ing to methods known to the person skilled in the art. The regenerated plant may be chimeric with respect to the incorporated foreign DNA. If the cells containing the foreign DNA develop into either micro- or macro-spores the integrated foreign DNA will be transmitted to a sexual progeny. If the cells containing the foreign DNA are somatic cells of the plant, non-chimeric transgenic plants are produced by conventional methods of vegetative propagation either in vivo, i.e. from buds or stem cuttings or in vitro following established procedures known in the art.

Thus, the present invention also relates to transgenic plants, parts of plants, plant tissue, plant seeds, plant embryos, plant seedlings, plant propagation material, plant harvest material, plant leaves and plant pollen, plant roots containing the above identified plants cell of the present invention. These plants or plant parts are characterised by, as a minimum, the presence of the heterologous transferred DNA construct of the present invention in the genome or, in cases where the transferred nucleic acid molecule is autologous to the transferred host cell are characterised by additional copies of the nucleic acid molecule of the present invention and/or a different location within the genome. Thus, the present invention also relates to plants, plant tissues, plant seeds, plant seedlings, plant embryos, propagation material, harvest material, leaves, pollen, roots, calluses, tassels etc. non-biologically transformed which possess stably or transiently integrated in the genome of the cells, for instance in the cell

nucleus, plastids or mitochondria a heterologous and/or autologous nucleic acid sequence containing (a) a coding sequence of the present invention or (b) a regulatory element of the present invention recognised by the polymerases of the cells of the said plant and, in a preferred embodiment, being operably linked in sense or antisense orientation to in case of (a) at least one regulatory element or in case of (b) a coding sequence of a gene of interest. The teaching of the present invention is therefore applicable to any plant, plant genus or plant species wherein the regulatory elements mentioned above are recognised by the polymerases of the cell. Thus, the present invention provides plants of many species, genuses, families, orders and classes that are ably to recognise these regulatory elements of the present invention or derivatives or parts thereof.

Any plant is considered, in particular plants of economic interest for example plants grown for human or animal nutrition, plants grown for the content of useful secondary metabolites, plants grown for their content of fibres, trees and plants of ornamental interest. Examples which do not imply any limitation as to the scope of the present invention are corn, wheat, barley, rice, sorghum, sugarcane, sugarbeet, soybean, *Brassica*, sunflower, carrot, tobacco, lettuce, cucumber, tomato, potato, cotton, *Arabidopsis*, *Lolium*, *Festuca*, *Dactylis*, or poplar.

The present invention also relates to a process, in particular a microbiological process and/or techni-

cal process, for producing a plant or reproduction material of said plant, including an heterologous or autologous DNA construct of the present invention stably or transiently integrated therein, and capable of being expressed in said plants or reproduction material, which process comprises transforming cells or tissue of said plants with a DNA construct containing a nucleic acid molecule of the present invention, i.e. a regulatory element which is capable of causing the stable integration of the ZmMADS2 coding sequence in said cell or tissue and enabling the sense or antisense expression of a ZmMADS2 coding sequence or part thereof in said plant cell or tissue, regenerating plants or reproduction material of said plant or both from the plant cell or tissue transformed with said DNA construct and, optionally, biologically replicating said last mentioned plants or reproduction material or both.

Finally, the present invention relates to a method for isolating or cloning pollen and/or root specific gene and/or pollen and/or root specific regulatory elements, such as promoters, or MADS box genes whereby a nucleic acid sequence of the present invention is used to screen nucleic acid sequences derived from any source, such as genomic or cDNA libraries derived from plants, in particular monocotyledonous plants. The nucleic acid sequences of the present invention thereby provide a means of isolating related regulatory sequences of other plant species which confer pollen or root specificity to genes of interest operably linked to them.

Further preferred embodiments of the present invention are mentioned in the claims.

The invention may be more fully understood from the following detailed sequence descriptions which are part of the present teaching. The SEQ ID Nos. 1 to 40 are incorporated in the present invention. The positions indicated below refer to the sequence numbering of SEQ ID No. 12 in 5' to 3' direction.

SEQ ID No. 1 represents the complete cDNA-sequence of the ZmMADS2 (*Zea mays* MADS-box) gene.

SEQ ID No. 2 represents the amino acid sequence of the ZmMADS2 protein.

SEQ ID No. 3 to 9 are partial DNA sequences of the ZmMADS2 gene of *Zea mays* and represent promoters and promoter fragments of various lengths.

SEQ ID No. 3: The 1502 bp (base pair) sequence spans the region from and including position 1 towards the 3' end, up to and including position 1502.

SEQ ID No. 4: The 502 bp (base pair) sequence spans the region from and including position 1 towards the 3' end, up to and including position 502.

SEQ ID No. 5: The 500 bp (base pair) sequence spans the region from and including position 503 towards the 3' end, up to and including position 1002.

SEQ ID No. 6: The 1000 bp (base pair) sequence spans the region from and including position 503 towards the 3' end, up to and including position 1502.

SEQ ID No. 7: The 740 bp (base pair) sequence spans the region from and including position 763 towards the 3' end, up to and including position 1502.

SEQ ID No. 8: The 500 bp (base pair) sequence spans the region from and including position 1003 towards the 3' end, up to and including position 1502.

SEQ ID No. 9: The 410 bp (base pair) sequence spans the region from and including position 1093 towards the 3' end, up to and including position 1502.

SEQ ID No. 10 is a partial DNA sequence of the ZmMADS2 gene of Zea mays. The 310 bp sequence spans the region from and excluding the translation initiation codon ATG at position 1813 to 1815 towards the 5' end up to the initiation site of transcription at position 1503 (5'-AAACG-3'). This sequence is the 5' untranslated region (5' UTR).

SEQ ID No. 11 is also a partial DNA sequence of the ZmMADS2 gene and represents a 3' regulation element of 251 bp from and excluding the translation stop codon TGA at position 4379 to 4381 to position 5031.

SEQ ID No. 12 is the genomic DNA sequence of the ZmMADS2 gene of Zea mays comprising 5031 nucleotides and encompassing the coding region of the

gene, an 5' untranslated region of 310 bp [5' UTR: position 1503-1812] between and excluding the translation initiation codon ATG (position 1813 - 1815) and including the nucleotide at position 1503 (also depicted in SEQ ID No. 10) and the regulatory elements identified in SEQ ID Nos 3 to 9 and 11.

SEQ ID No. 13 to 40 represent primers used for cloning and detecting nucleic acid sequence of the present invention and/or transcripts expressed thereby.

The invention is further illustrated by way of example and the following drawings.

Figure 1 shows a northern blot analysis for various tissues of Zea mays. A signal was obtained exclusively in mature pollen.

Figure 2 shows RT-PCR analyses of various maize tissues. A signal could be obtained in mature pollen and root tips (lower band; 350 bp).

Figure 3 shows PCR analyses of various reproductive tissues, in particular of cDNA libraries of egg cells, in vitro zygotes 18 hours after fertilisation, mature pollen and leaves (from seedlings). A signal was obtained only in mature pollen.

Figure 4 shows RT-PCR analyses of ZmMADS2 expression during microgametogenesis of maize. A signal was obtained in mature pollen after anthesis.

Figure 5 shows ZmMADS2 transcript in growing pollen tubes of maize.

Figure 6 shows transient transformation assays of *N. tabacum*.

Example 1: Cloning of the ZmMADS2 cDNA sequence

Plant material and pollen isolation

Tissues were isolated from *Zea mays* L. inbred line A188 (Green and Philips, 1975) cultivated in a greenhouse. Embryos from kernels (12 dap and mature) were isolated under sterile conditions. Seedlings were germinated under sterile conditions in the dark and were dissected into cotyledons, roots tips and scutella. For isolation of pollen before anthesis, tassels were divided into upper (mature stage) and lower parts (immature stage). Pollen was isolated as described by Mordhorst et al. (1993) and different developmental stages were separated via a discontinuous Percoll gradient (20%, 30% and 40% Percoll in 0.4 M mannitol). Centrifugation was performed for 10 min at 20° C and 226 x g in a swing out rotor with slow acceleration and deceleration. Developmental stages of pollen were monitored microscopically and by DAPI staining (see below).

RNA isolation and construction of cDNA libraries

Total RNA was isolated from various tissues with TRIzol (GibcoBRL). Seasand was added for the mac-

eration of pollen. Total RNA was isolated from mature pollen for the construction of a cDNA library using the protocol described by Stirn et al. (1995). The library was constructed from 5 µg poly(A)⁺ RNA as outlined by Dresselhaus et al. (1996b) using the Uni-ZAP XR lambda vector (Stratagene). Total RNA from leaf material of 10 day old seedlings was isolated as described by Logemann et al. (1987) and a cDNA library was generated from 2 µg poly(A)⁺ RNA (seedlings library).

Screening of cDNA library with maize MADS box probes

The highly conserved MADS box of different maize MADS box genes was amplified from the maize genome by PCR and served as probes for the plaque screening of a cDNA library of mature maize pollen.

Genomic DNA from leaf material was isolated as outlined by Dellaporta et al. (1983) and served as template for the synthesis of different MADS box probes. Gene specific primers with the nucleotide sequence specified in SEQ ID No. 13 to 24 were used to specifically amplify the MADS box region of maize MADS box genes:

ZMM1 (5'-ATGGGGAGGGGAAGGATTGA-3', SEQ ID No. 13;
5'-CTGTTGTTGGCGTACTCGTAG-3', SEQ ID No. 14),
ZEM2/3/ZAG 4 (5'-AGGGGCAAGATCGACATCAAG-3', SEQ ID No. 15;
5'-GG/TCGT/AACTCGTAGAGGCAG-3', SEQ ID No. 16),
ZAG3/5 (5'-ATGGGGAGGGGACGA/CGTTGA-3', SEQ ID No. 17;

5'-GCTGCCGAACCTCGTAGAGCT-3'; SEQ ID No. 18), ZAP1 (5'-GTTGTTGGCGTACTCGTAGAG-3', SEQ ID No. 19; 5'-GGGCGCAAGGTACAGCTGAA-3', SEQ ID No. 20), ZAG1 (5'-GTTGTTGGCGTACTCGTAGAG-3', SEQ ID No. 21; 5'-AAGGGCAAGACTGAGATCAAG-3, SEQ ID No. 22) and ZAG2 (5'-CACTTGAACCTTTACGCTTAT-3', SEQ ID No. 23; 5'-GACAATCTTGACACATGTATGAA-3', SEQ ID No. 24); amplification of the MADS box and flanking genomic regions: PCR amplification was performed with 200 ng genomic DNA in a standard reaction mixture: 250 nM primer, 2 mM MgCl₂, 400 μM dNTPs and 1.25 U Taq DNA polymerase (GibcoBRL) in PCR buffer (50mM KCl, 20 mM Tris-HCl, pH 8.4). Hot start PCRs were performed with the following profile: 5 min 95° C, 3 min 75° C (addition of Taq-DNA polymerase) followed by 30 cycles with 1 min 96° C, 1 min 62° C (ZEM, ZMM, ZAG3) or 60° C (ZAG1, ZAP1) and 3 min 72° C. A final extension was performed for 5 min at 72° C. PCR products were separated on low melting agarose gels (NuSieve GTG, BIOzym) and isolated gel fragments containing the DNAs were digested with Gelase (BIOzym). Probes were labelled with [³²P]-dCTP (6000 Ci/mmol), Amersham) using the Prime-it II random primer labelling Kit (Stratagene) and purified with NucTrap columns (Stratagene). Approximately 22.000 phages from the pollen library were plated per 15 cm plate and transferred by Hybond-N membranes (Amersham) as double plaque lifts according to Sambrook et al. (1989). Prehybridisation was performed with 50 μg/ml salmon sperm DNA in hybridisation buffer (5xSSPE, 5x Denharts, 0.5% SDS) for 5 h at 55° C. Filters were hybridised with a cocktail of the different MADS box probes in a final concentration of 650.000 cpm for each probe/ml

hybridisation buffer. After hybridisation overnight at 55° C, filters were washed three times for 15 min with 5x SSPE/0,1% SDS and exposed to X-Omat AR films (Amersham) using intensifier screens at -70° C. Putative positive lambda phages were isolated and cDNAs excised according to the manufacturer (ZAP-cDNA Synthesis Kit, Stratagene).

Thus, approximately 250.000 phages were screened with the MADS box probes at medium stringent conditions to permit hybridisation to less homologous sequences. Thirteen putative positive signals were analysed further and two cDNAs with high homology to MADS box proteins were isolated, designated ZmMADS1 and ZmMADS2.

5' RACE

ZmMADS2 cDNA isolated from the pollen cDNA library was incomplete at its 5' end and a PCR approach was chosen to isolate the full length cDNA from the pollen library. For amplification of the missing 5' end, a gene specific RACE-primer (rapid amplification of cDNA ends, 5'-CGATTCAAATCGTGAATCTCAT-3'; SEQ ID No. 25) combined with a pBluescript vector primer (5'-CCCCCGGGCTGCAGGAATTC-3'; SEQ ID No. 26-primer) was used in a standard PCR reaction with the following profile: 3 min 96° C followed by 33 cycles: 20 sec 96° C, 30 sec 62°, 2 min 72° C and a final extension for 5 min at 72° C. PCR products were cloned and sequenced. Full length cDNA was amplified from 3 µl pollen cDNA library with primers specific for the ZmMADS2 3'-end and 5'-end (5'-TTTTAGCAACATCTGCACCATT-3'; SEQ ID No.

27; 5'-TCTCGGCTAGCTTCCTCCT-3'; SEQ ID No. 28) using a mixture of 2 U Pfu DNA polymerase and 1 U TaqPlus Long DNA polymerase (Stratagene) for a first amplification round of 25 cycles in a standard PCR reaction with Pfu reaction buffer (10 mM KCl, 10 mM (NH₄)₂SO₄, 2 mM MgSO₄, 0.1% TritonX100, 20 mM Tris-HCl, pH 8.75). 1/10 of purified PCR products served as template in a subsequent second amplification round with Pfu DNA polymerase (25 cycles). PCR profiles for both amplification rounds were as follows: 3 min 96° C, 5 min 75° C followed by 25 cycles with 20 sec 96° C, 30 sec 58° C, 90 sec 72° C and a final extension for 5 min at 72° C. PCR products were cloned and sequenced.

DNA Sequencing and sequence analysis

Sequencing of cDNAs was performed with the ABI PRISM Dye Terminator Kit with TaqFS DNA polymerase (PE Applied Biosystems) according to the manufacturers protocol, except that 800 ng of template DNA and 5 pmole vector primers were used. Sequence analyses were performed with DNASIS 1.1 software program package (HITACHI).

The cDNA of ZmMADS2 is 1268 bp in length with an open reading frame of 240 amino acids. The 3'-UTR is 251 bp (3' UTR: position 3976 - 4226) and the 5'-UTR is 310 bp (5' UTR: position 1503 - 1812) in length, calculated from the putative start point of transcription at position 1503 as analysed by primer extension (data not shown) to and excluding the ATG at position 1813-1815.

The sequence of the full-length cDNA is given in SEQ ID No. 1. The amino acid sequence of ZmMADS2 is given in SEQ ID No. 2. ZmMADS2 contains a MADS-box at the N-terminal end consisting of 57 (including the start methionine) amino acids. The MADS-box is followed by a linker region of 32 amino acid and a K-box comprising 67 amino acids. A putative bipartite nuclear localisation signal is located in the MADS box of ZmMADS2. A bipartite signal motive (RR-(X)₁₂-KRR) is comprised of two basic amino acids, a spacer of twelve variable amino acids and a basic cluster in which three out of five amino acids are basic. Putative N-glycosylation sites (N-X-T/S) and several putative phosphorylation sites for protein kinase C (ST-X-RK) and Casein Kinase II (ST-X₂-DE) have been found. Homology searches revealed that ZmMADS2 belongs to the AGL 17 family of MADS box proteins (Zachgo et al., 1997).

Example 2: Cloning of the genomic ZmMADS2 clone

The cDNA isolated in Example 1 was used to clone the genomic clone as follows:

Genomic DNA from leaf material of Zea mays line A188 was used to construct libraries with the "Universal Genome Walker Kit" (Clontech, Palo Alto). These libraries served as templates for the synthesis of four genomic fragments. The amplification was carried out with nested adapter primers delivered by the manufacturer and nested gene specific primers (Tnorf 1: 5'- CGGCCTATAGCTAGCTCTCTTGACCCT-3', SEQ ID No. 29;

Front1: 5'-AGGGTCAAGAAGAGAGCTAGCTATAGG-3', SEQ ID No. 30; Tnorf2: 5'-GCTAAGGAGCGAGAGGTTGTGGTTGTGG-3', SEQ ID No. 31; Front2: 5'-CCACAACCACAAACCTCTCGCTCCTTA-3', SEQ ID No. 32 according to the manufacturer's instructions.

PCR amplification of a fragment containing intron 1 and the first 247 bp of intron 2 was performed following the "Genome Walker" protocol. For amplification of the promoter fragments, 5% DMSO was added to the reaction mixture. Synthesis of the sequences of introns 3 to 6 was performed in a standard PCR reaction mixture with 200 ng genomic DNA template and the following gene specific primers: Ex3_ober: 5'-TCGGCAGTTGACGGGAGAT-3', SEQ ID No. 33, In3'_low: 5'-TTAGCAACTCATTATAGCAC-3', SEQ ID No. 34, Pe_for: 5'-AGAAACCAGAGATGTTCCAG-3', SEQ ID No. 35, Pe_rev: 5'-CAACATCTGCACCATTGGAA-3', SEQ ID No. 36, Intron2_1: 5'-GTCACATCAACAAGCGTCCTACC-3', SEQ ID No. 37, Intron3: 5'-TAAGAGCATCAACAGTAGTAACG-3', SEQ ID No. 38.

All fragments were cloned with the Zero blunt TOPO PCR Cloning Kit (Invitrogen, Leek) and sequenced at SEQLAB (Göttingen), except for fragments containing Intron 2. Amplification of Intron 2 with different primer pairs (Intron2_1 and Pe_rev, Intron2_1 and In3'_low, Intron 2_1 and Intron3, Front2 and Intron3) using different DNA polymerases (PfuTurbo, Stratagene; PlatinumTaq, GibcoBRL; AdvantageTth, Clontech) showed that this Intron is about 11 kb in length.

The nucleotide sequence of the genomic clone as depicted in SEQ ID No. 12 comprises 5031 nucleotides. Intron No. 2 has not been fully sequenced. Depending upon the start of transcription, the promoter comprises nucleotides 1 to 1502, since a transcription start point has been determined at position 1503 5'-AACGC-3'. From position 1503 to the ATG initiation translation codon at position 1813 to 1815 is the 5' untranslated region. Exons and consequently also introns are indicated in SEQ ID No. 12. In intron II between position 3203 and 3204 there remains some sequence ambiguity. Further nucleotides are present in intron II at this position (circa 11 kb).

The translation termination codon TGA is located at position 4778 to 4780 with a transcription termination signal between and including position 4781 and 5031. A putative polyadenylation site TATAA is located at position 4888 to 4893.

The promoter from position 1 to position 1502 depicted in SEQ ID No. 3 may be divided into functional promoter fragments with sequences as indicated in SEQ ID No. 4 to 9. SEQ ID No. 4 comprises at position 17 to 27 an induction element responsive to dryness, abscisic acid and/or coldness and at position 151 to 155 a root-specific element. The promoter fragment of SEQ ID No. 5 comprises at positions 532 to position 536, position 542 to 546, positions 634 to 638 and position 705 to 709 a root-specific element, at position 654 to 658 an element responsible for induction by sugar deple-

tion and at position 684 to 689 an element responsive to dryness, abscisic acid and/or coldness.

The promoter set out in SEQ ID No. 8 comprises an element responsible for induction by dryness, abscisic acid and/or coldness located at position 1019 to 1025 and 1369 to 1375, a root-specific element at position 1083 to 1087 and a sucrose responsive element at position 1192 to 1200, as well as an element responsible for induction by sugar depletion at position 1259 to 1263. The element conferring root-specificity at position 1083 to 1087 is particularly important due to its root-specificity. Deletion of this element in the promoter fragments and promoters of the present invention provides a promoter being exclusively expressed in pollen.

In the 5' UTR depicted in SEQ ID No. 10, an element responsible for induction by sugar depletion is located at position 1701 to 1708.

Further elements conferring root-specificity are located at the following positions: 2281 to 2286, 2518 to 2522, 2575 to 2579, 2802 to 2807, 3014 to 3018, 4295 to 4299, 4506 to 4510, 4700 to 4704, 4888 to 4892.

Further elements responsible for induction by dryness, abscisic acid and/or coldness are located at the following positions: 2014 to 2019, 2055 to 2060, 2072 to 2077, 2392 to 2397, 2584 to 2589, 2679 to 2683, 2842 to 2846, 2863 to 2869, 3112 to 3117, 3231 to 3236, 4067 to 4072, 4228 to 4233,

4042 to 4246, 4344 to 4350, 4479 to 4484, 4486 to
4492.

Example 3: Northern blot and PCR analyses

Ten μ g of total RNA extracted from various tissues were separated on denaturating agarose gels and transferred to Hybond N' membranes (Amersham) by capillary blotting with 10x SSC overnight. The RNA was fixed to the membrane by UV crosslinking with 300 mJoule in a Stratalinker 1800 (Stratagene). Gene specific probes were amplified from plasmids containing ZmMADS2 cDNAs with primers specific for the 3'-ends of ZmMADS2 (5'-AGAAACCAGAGATGTTCCAG-3'; SEQ ID No. 39, 5'-CAACATCTGCACCATTGAA-3'; SEQ ID No. 40) according to the profile described for the 5'RACE. Probes were labelled as described above. Prehybridisation was performed for 5 h at 65° C with 100 μ g/ml salmon DNA in hybridisation buffer (7% SDS, 1 mM EDTA, 0.5 M NaH₂PO₄, pH 7.2). Probes were added in a final concentration of 10⁶ cpm/ml hybridisation buffer and hybridised overnight at 65° C. Filters were washed with decreasing concentrations of SSC with a final wash of 0.2xSSC/0.1% SDS for 15 min at 65° C and were exposed to X-Omat AR films (Amersham) at -70° C using intensifier screens. Fig. 1 shows the spatial and temporal expression of ZmMADS2 (dag: days after germination, dap: days after pollination). It is shown that the ZmMADS2 coding sequence is expressed in pollen, in particular in mature pollen after dehiscence indicating an essential role in pollen development and pollen function, in particular pollen tube growth.

RT-PCR analyses of tissues for ZmMADS2 expression were performed with 1 µg of total RNA (Figure 2). RNA was reverse transcribed with 50U SuperScriptII reverse transcriptase (GibcoBRL) for 75 min at 50° C with the ZmMADS2 specific primer (see above) in reaction buffer (75 mM KCl, 3 mM MgCl₂, 50 mM Tris-HCl, pH 8.3) with 0.4 mM dNTPs, 10 mM DTT and 10 U RNasin (Promega). The complete RT-PCR reaction was reused as template in the subsequent standard PCR reaction with ZmMADS2 specific primers (see above). PCR products were separated on an agarose gel blotted and hybridised to radiolabelled ZmMADS2 specific probe. Figure 2 shows ZmMADS2 expression in mature pollen and root tips (lower band).

For quantitative RT-PCR analysis of ZmMADS2 expression in pollen, RNA samples were treated with DNaseI according to the manufacturer's protocol (GibcoBRL). 250 ng and 500 ng, respectively, of each RNA probe were reverse transcribed and a standard PCR was performed with 27 cycles (ZmMADS2) as described above (Figure 4). cDNA libraries of maize egg cells, in vitro zygotes (Dresselhaus et al., 1994 and 1996a), pollen and seedlings were analysed for the presence of ZmMADS2 cDNAs with gene-specific primers using 3 µg of cDNA libraries as templates in a standard PCR as described above. To confirm specificity of PCR products, gels were blotted and hybridised to radiolabelled gene-specific probes (Fig. 3). Fig. 3 shows that among various reproductive cells ZmMADS2 is only expressed in pollen.

Thus, it was shown that ZmMADS2 is most abundant in mature pollen after dehiscence. At that stage pollen are dehydrated and metabolically inactive. The onset of ZmMADS2 expression therefore takes place after the pollen reaches maturity, but before dehydration inhibits transcriptional activity (Figure 4). A faint signal was detected in root tips (Fig. 2) which is confirmed by RT-PCR analyses as shown in Fig. 2. The lower band represents amplified ZmMADS2 cDNA which is detectable in mature pollen and root tips. The upper band represents genomic DNA.

In order to localise ZmMADS2 transcripts whole mount RNA in situ hybridisation experiments were performed.

Example 4

Whole mount in situ hybridisation and DAPI staining

Pollen was germinated in liquid medium as described by Torres et al. (1995). Germination rates were determined microscopically and only samples with high germination rates were used for the experiments. All media were prepared with DEPC-treated deionized water. Germinated pollen were fixed for 2 to 3 h in 4% paraformaldehyd in germination medium at room temperature. Fixation and all following steps were performed in Eppendorf tubes and pollen was precipitated by centrifugation for 1 min at 1500 rpm at room temperature for the exchange of medium. Pollen was dehydrated by passage through a series of 3:7, 1:1, 7:3 ME:PP solutions at room

temperature for 30 min each (ME: 90% methanol and 10% 0.5 M EGTA; PP: 4% paraformaldehyde in PBS (phosphate buffered saline, 130 mM NaCl, 10 mM Na₂PO₄, pH 7.4)). Pollen was incubated for 10 min in 100% ethanol and for 10 min in 100% xylol before rehydration through an ethanol series was performed (85%, 70%, 55%, and 35% ethanol in water). Proteinase K (Boehringer) treatment was carried out for 35 min at 37° C with 1 µg/ml proteinase in 100 mM Tris-HCl (pH 7.5) with 50 mM EDTA. Reactions were stopped with 2 mg/ml glycine in PBS (2 min) and pollen was washed in PBS (2 min). Pollen were re-fixed for 20 min at room temperature with 4% paraformaldehyde in PBS and were washed twice for 5 min with PBS. Prehybridisation was performed with 100 µg/ml tRNA and 100 µg/ml mRNA (Boehringer) in 1 ml hybridisation buffer (6x SSC, 0.1% SDS, 50% Formamid) at 42° C for 3 h with gentle shaking. After prehybridisation, 800 µl of the buffer was removed and the remaining 200 µl were distributed to two 500 µl Eppendorf tubes. Probes were added to a final concentration of 500 ng/ml. DIG-labelled RNA probes were synthesised from the 3' end of ZmMADS2 cloned into pGEM-T vector (Promega) with Sp6 and T7 RNA Polymerase (Boehringer) according to the manufacturer's protocol. Hybridisation was carried out at 42°C overnight without shaking. The samples were washed for 10 min at 42° C with 2x SSC, 0.1% SDS with gentle shaking before RNase treatment with 40 µg/ml RNaseA (Boehringer) was performed for 30 min at 37° C in 150 µl RNase buffer (10 mM Tris-HCl (pH 8), 0.5 M NaCl, 1 mM EDTA). Samples were washed three times for 10 min in 1x SSC, 0.1% SDS and 0.5x SSC, 0.1% SDS at 42° C and equilibrated for 5 min

in B1 buffer (Boehringer). Blocking was performed in B1 buffer with 1% blocking reagent (Boehringer) and 0.3% TritonX100 for 1h at room temperature. Detection was performed with NBT/BCIP according to the manufacturer's protocol (Boehringer). The reactions were monitored microscopically and stopped by adding buffers at neutral pH (Tris-HCl or B1-buffer). Reactions could be restarted by retransferring the probes into reaction mixture.

DAPI staining was performed with pollen fixed in ethanol:acetic acid (3:1) for 1 h at room temperature and dehydrated through an ethanol series (75%, 55%, 35% ethanol in water). Pollen was transferred to Tris-HCl buffer (pH 7.5) with 200 ng/ml DAPI and staining was monitored under UV light (359/441 nm).

The *in situ* hybridisation experiments with ZmMADS2 antisense probes showed transcripts in the emerging pollen tube 5 min after germination (Figure 5a). During pollen tube growth, transcripts are translocated into pollen tubes displaying a gradient with increasing concentration toward the tip (Figure 5c). The translocation of ZmMADS2 transcripts into the pollen tube precedes migration of the vegetative nucleus in the sperm cells into pollen tubes by more than two hours. Possibly, ZmMADS2 is required only later in the germination process. ZmMADS2 transcript levels in microspores are very low (Figure 4). Maximal expression in mature pollen after dehiscence and localisation of transcripts in pollen tubes clearly classifies ZmMADS2 as a late pollen gene. The above findings clearly imply a function for ZmMADS2 during pollen tube

growth, in particular a function as a transcription factor, i.e. in regulating the transcription of one or more target genes.

Example 5: Transient transformation assays

Pollen transformation was carried out according a modified protocol of Hamilton et al. (1992). Plasmids containing the ZmMADS2 promoter fragments fused to the luciferase reporter gene and the NOS terminator were used to coat gold particles (0,3 μ m - 3 μ m) which were then bombarded into pollen grains of *Nicotiana tabacum* (SR1) using a helium driven ballistic device (PDS-1000/He Biostatic Particle Delivery System, Bio-Rad). The plasmids contained fragments of SEQ ID No. 3, Pos. 1-1538 (pDNS5), 455-1538 (pDNS6), 1006-1538 (pDNS7), 1175-1538 (pDNS8) or no promoter (negative control).

Pollen were spread out with a brush on sterile plastic petri dishes (30 mm x 15 mm) containing pollen medium (0,01% H_3BO_3 , 10 mM $CaCl_2$, 0,05 mM KH_2PO_4 , 0,1% yeast extract and 10% sucrose) with 0,6% agarose. Bombardments were done 5-20 min after spreading of the pollen. For bombardments done with leaf tissue, tips of young leaves were cut and placed on petri dishes containing sterile water with 0,6% agarose. The plates were kept at room temperature for 12-16 h prior to measurement of luciferase activity.

Measuring of luciferase activity were carried out with the Luciferase Assay Kit (Promega) following the protocol of the manufacturer (Technical Bulle-

tin #TB101) except for the preparation of the tissues. Germinated pollen, along with the agarose medium, were transferred to ice cold tubes followed by addition of 0,2 ml of cold lysis buffer and ground with a manual homogenizer. Leaf tips were transferred into ice cold tubes and homogenized in a Retschmill (MM 2000). Thereto 0,2 ml of lysis buffer was added. All samples were centrifuged at 10,000 x g for 5 min. at 4°C. The supernatants were assayed for luciferase activity in a Luminometer (Lumat LB 9501, Berthold). Corrections were made for the volume of liquid remaining in the agarose pellets. Total numbers of assayed samples per construct were 30 (pollen) or 5 (leaves).

References cited:

An et al., EMBO J. 4 (1985), 277-287.

Bate, N. Spur, C. Foster, G. and Twell, D., Plant J. 10 (1996), 613-623.

Brettschneider et al., Theor. Appl. Genet. 94 (1997), 737-748

Bytebier et al., Proc. Natl. Acad. Sci. USA 84 (1987), 5345-5349.

Chan et al., Plant Mol. Biol. 22 (1993), 491-506.

Curie, C. and McCormick, S., Plant Cell, 9 (1997), 2025-2036.

Dellaporta, S.L., Wood, J. and Hicks, J.B., Plant Mol. Biol. Rep. 1 (1983), 19-21.

Dresselhaus, T., Lörz, H. and Kranz, E., Plant J. 5, (1994), 605-610.

Dresselhaus, T., Hagel, C., Lörz, H. and Kranz, E., Plant Mol. Biol. 31 (1996a), 23-34.

Dresselhaus, T., Barcelo, P., Hagel, C., Lörz, H. and Humbeck, K., Plant Mol. Biol. 30 (1996b), 1021-1033.

Estruch, J.J., Kadwell, S., Merlin, E. and Crossland L., Proc. Natl. Acad. Sci. USA, 91 (1994), 8837-8841.

Gould et al., Plant Physiol. 95 (1991), 426-434.

Green, C.E. and Philips, R.L., Crop Science, 15 (1975), 417-421.

Hamilton et al., Plant Mol. Biol., 18 (1992), 211-218.

Hiei et al., Plant J. 6 (1994), 271-282.

Hoekema: The Binary Plant Vector System, Offset-drukkerij Kanters B.V., Albllasserdam (1985), Chapter V.

Holsters et al., Mol. Gen. Genet. 163 (1978), 181-187.

Ishida et al., Nature Biotechnology 14 (1996), 745-750

Jähne et al., Eurphytica 85 (1995), 35-44.

Lee, H.-S., Karunanananda, B., McCubbin, A., Gilroy, A. and Kao, T.H., Plant J. 9 (1996), 613-624.

Li, H.-Y., and Gray, J.E., Plant Mol. Biol. 33 (1997), 653-665.

Lit et al., Plant Mol. Biol. 20 (1992), 1037-1048.

Logemann, J., Schell, J. and Willmitzer, L., Anal. Biochem. 163 (1987), 16-20.

Maheshwari et al., Critical Reviews in Plant Science 14 (2) (1995), 149-178.

Mooney et al., Plant Cell Tiss. & Org. Cult. 25 (1991), 209-218.

Mordhorst, A.P. and Loerz, H. J., Plant. Physiol. 142 (1993), 485-492.

Mu, J.H., Lee, H.-S. and Kao, T.H., Plant Cell, 6 (1994), 709-721.

Muschietti, J., Eyal, Y. and McCormick, S., Plant Cell, 10 (1998), 319-330.

Op den Camp, R.G.L. and Kuhlemeier, C., Nucleic Acid Res. 26 (1998), 2058-2062.

Raineri et al., Bio/Technology 8 (1990), 33-38.

Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd. Edition, Cold Spring Harbor Press, NY, (1989).

Stirn, S., Mordhorst, A.P., Fuchs, S. and Lörz, H., Plant Sci. 106 (1995), 195-206.

Taylor, L.P. and Hepler, P.K., Annu Rev. Plant Phys. Plant Mol. Biol. 48 (1997), 461-491.

Torres et al., Plant J. 8 (1995), 317-321.

Wilson, C., Voronin, V., Touraev, A., Vicente, O.
and Heberle-Bors, E., Plant Cell, 9 (1997), 2093-
2100.

Zachgo, S., Saedler, H., Schwarz-Sommer, S.S.,
Plant J. 11 (1997), 1043-1050.

Claims

1. A nucleic acid molecule for use in cloning and expressing a pollen specific nucleic acid sequence in a plant which is selected from the group consisting of
 - (a) the nucleic acid sequence defined in any one of SEQ ID No. 1 or 12, or a complementary strand thereof,
 - (b) a nucleic acid sequence encoding a protein or peptide with the amino acid sequence defined in SEQ ID No. 2 or a complementary strand thereof,
 - (c) a nucleic acid sequence which hybridises to the nucleic acid sequence defined a) or b), or a complementary strand thereof and
 - (d) a nucleic acid sequence which is degenerate as a result of the genetic code to the nucleic acid sequence defined in a), b), c), or a complementary strand thereof,
 - (e) alleles or derivatives of the nucleic acid sequence defined in (a), (b), (c), (d), or a complementary strand thereof.

2. A nucleic acid molecule for use in cloning and expressing a pollen specific nucleic acid sequence in a plant which is selected from the group consisting of

(a) the nucleic acid sequence set out in SEQ ID NO. 3 to 11, or a complementary strand thereof.

(b) a nucleic acid sequence which hybridizes to the nucleic acid sequence defined in (a), or a complementary strand thereof and

(c) alleles or derivatives of the nucleic acid sequence defined in (a) or (b), or a complementary strand thereof.

3. The nucleic acid molecule of claim 1 or 2, which is derived from maize.

4. The nucleic acid molecule of claim 1, 2 or which is a DNA, cDNA or RNA molecule.

5. A vector comprising the nucleic acid molecule of any one of claims 1 to 4.

6. The vector of claim 5, which is a bacterial or viral vector.

7. The vector of any one of claims 5 or 6, wherein the nucleic acid molecule of any one of claims 1, 3 or 4 is operably linked to at least one regulating element, in particular in antisense orientation.

8. The vector of any one of claims 5 to 7, wherein the regulatory element is a 5' or 3' element.

9. The vector of claim 8, wherein the 5' regulatory element is a promoter, in particular the CaMV 35S promoter.

10. The vector of claim 8 or 9, wherein the 3' regulatory element is a termination and poly A addition sequence, in particular from the NOS gene of *Agrobacterium tumefaciens*.

11. The vector according to any one of claims 5 to 10, which furthermore contains T-DNA, in particular the left, the right or both T-DNA borders.

12. The vector according to claim 11, wherein the nucleic acid molecule, optionally in conjunction with at least one regulatory element, is located within the T-DNA or adjacent to it.

13. A host cell containing the vector of any one of claims 5 to 12 or a cell deriving therefrom.

14. The host cell of claim 13, which is a plant, yeast or bacterial cell, in particular a cell from a monocotyledonous or dicotyledonous plant or a cell deriving therefrom.

15. A cell culture, preferably a plant cell culture comprising a cell according to any one of claims 13 or 14.

16. A method of genetically modifying a cell by transforming a cell with a nucleic acid molecule of any one of claims 1 to 4 or a vector according to any one of claims 5 to 12, wherein the nucleic acid molecule of claims 1 to 4 contained in the vector is expressible in the cell.

17. The method of claim 16, wherein the cell is a plant, bacterial or yeast cell.

18. The method of claims 16 or 17, wherein the transformed cell is regenerated into a differentiated plant.

19. The method of any one of claims 16 to 18, wherein the cell is transformed by transfer of the nucleic acid molecule or vector from a bacterium to the cell.

20. The method of any one of claims 16 to 19, wherein the cell is transformed by direct uptake of nucleic acid sequences, by microinjection of nucleic acid sequences or particle bombardment.

21. A method for isolating a pollen specific gene from a plant, whereby a nucleic acid sequence of any one of claims 1 to 4 is used to screen nucleic acid sequences derived from the plant.

22. A plant comprising a host cell according to any one of claims 13 or 14 or produced according to a method according to any one of claims 16 to 21 or progeny thereof.

23. Propagation and harvest material, in particular seeds and plant tissue, comprising a host cell according to any one of claims 13 or 14 or derived from a plant according to claim 22.

24. A method for the production of a male sterile genetically modified plant with a distorted pollen development, wherein a plant cell is transformed with a nucleic acid molecule according to claims 1 to 4 or a vector according to claims 5 to 12 and the transformed cell is regenerated into a plant.

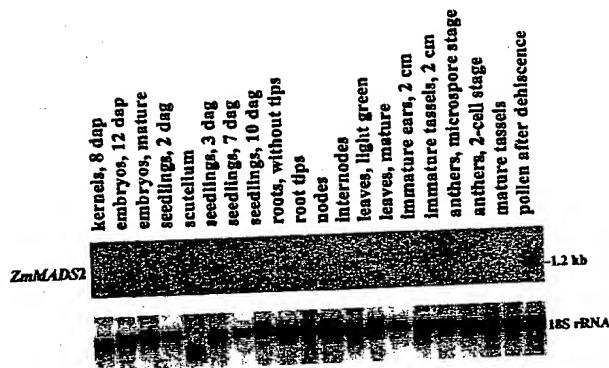


Figure 1

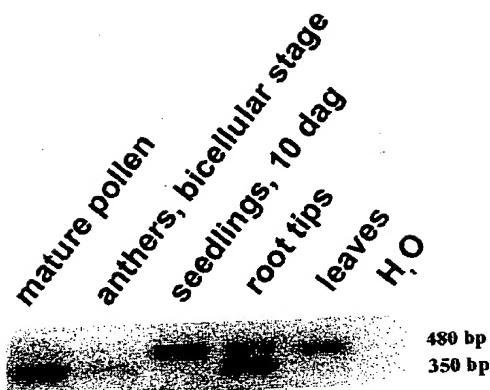


Figure 2

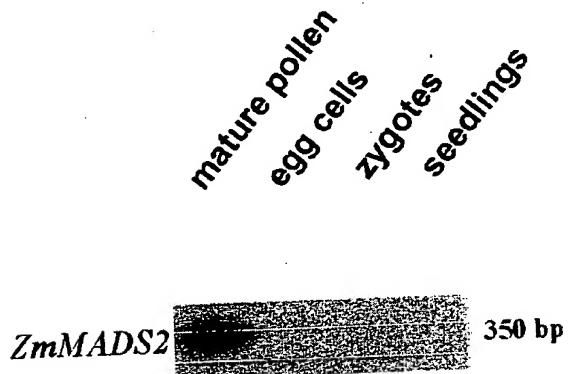


Figure 3

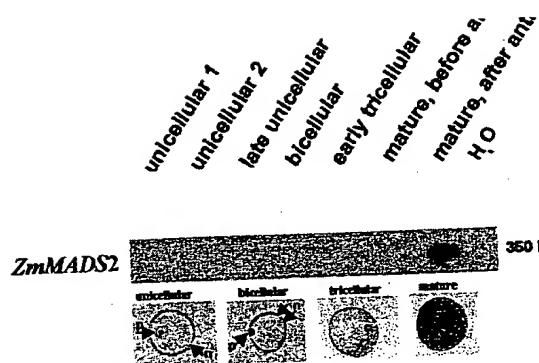


Figure 4

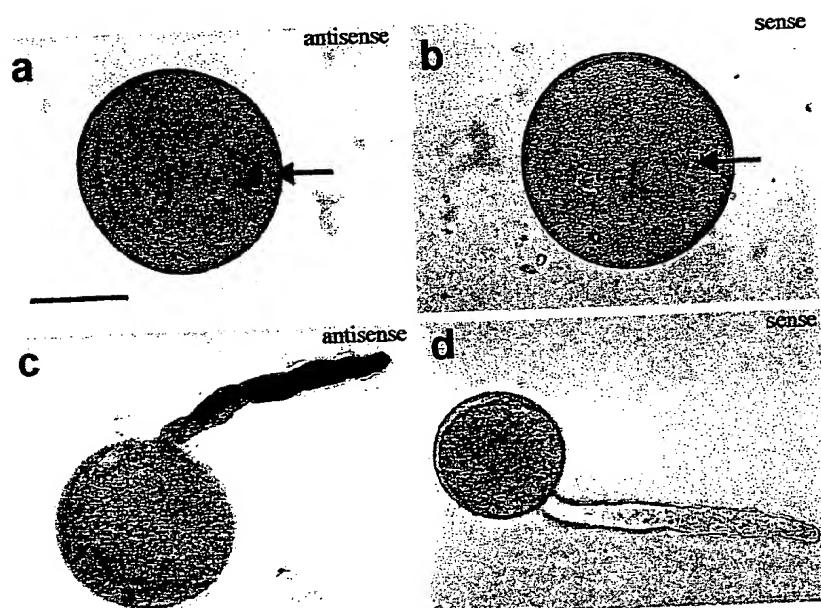


Figure 5

Transformation of *Nicotiana tabacum* - transient assays

Expression of luciferase in mature pollen and young leaves directed by deletion-constructs of the *ZmMADS2* - promoter

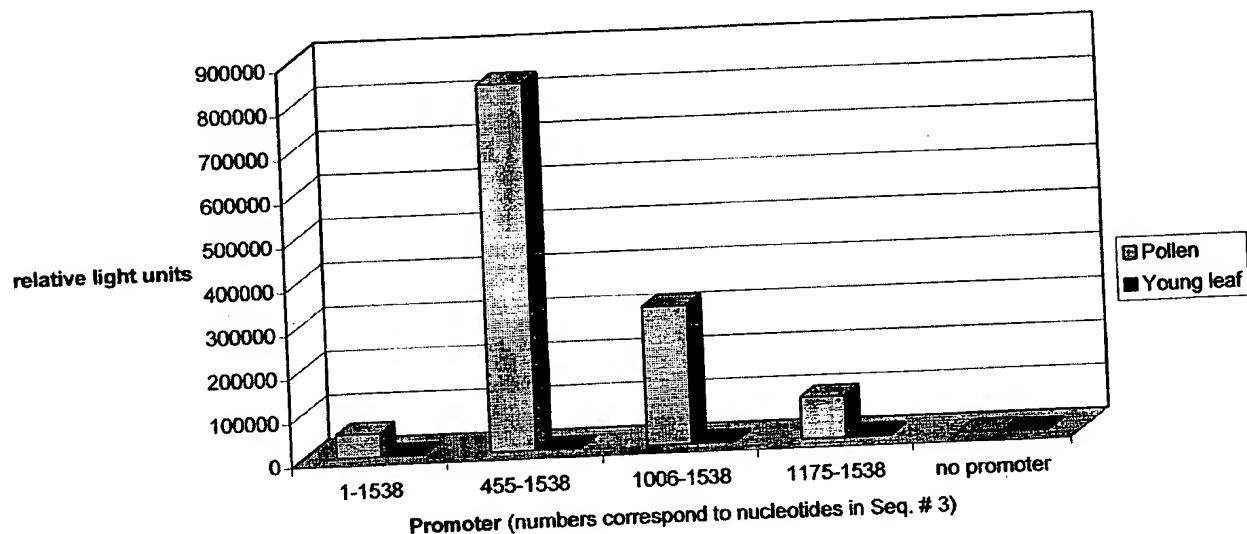


Figure 6

SEQUENCE LISTING

<110> Südwestdeutsche Saatzucht Dr. H. R. Späth

<120> Male sterile plants

<130> 23596 WO

<140>

<141>

<160> 40

<170> PatentIn Ver. 2.1

<210> 1

<211> 1268

<212> DNA

<213> Zea mays

<400> 1

tctcggctag cttecttcctc tctccctcctc ctctgtcttc ctctttccctt ctccccttggg 60
 aaacctgtcg cctttgagct ttcttttctcg agtgcgtccca ccagatctcc tcctcccttac 120
 ctctttggc acgttccggcg ggcgcgcgcgg agatagatag atccccccat cgtcgctgcgc 180
 gtcttgcctt ccgatcggag ggcacaacc acaacctctc gctctttagc gtgcgcgcgc 240
 gagccagggt caagaagaga gctagctata ggcggagat cgatggggag gggaaagatc 300
 gtgatecgca ggatcgataa ctccacgagc cggcagggtga ccttctccaa ggcgcggaaac 360
 gggatcttca agaaggccaa ggagctcgcc accctctgcg atgcggaggt cggcctcgctc 420
 atcttctcca gcacccggccg cctctacgag tactctagca ccagcatgaa atcagttata 480
 gatcggtagc gcaaggccaa ggaagagcag caagtctgcg caaatccaa ctcggagctt 540
 aagtttggc aaaggaggc agcaagctt agacaacaac tgcacaactt gcaagaaaat 600
 tatcggcagt tgacgggaga tgatctttt cggctgtatc tcaaaaactt gcagtccttg 660
 gagaatcaat tggaaacaag cctgcgtggg gtccgcgc当地 agaaggacca tctcttgata 720
 gatgagatc acgatttggaa tcgaaaggca agtttatttc accaagaaaa tacagactt 780
 tacaataaga tcaacctgtat tcgccaagaa aatgtgatgatc tacataaaaaa gatctatgag 840
 actgaaggac caagtggagt taatcgggag tcacggactc cattcaactt tgcatgtat 900
 gaaaccagag atgttccagt gcaacttgaa ctcagcacac tgccacaaca aaataacatt 960
 gagccatcata ctgctccaa gctaggattt caattaaatc catgaagaag agtaaaactg 1020
 ccgtctttagt atgtcgtatc aaacttatttta ttgtgaagag atgataactca gagaagaca 1080
 tattttgtggc agggagatc gagatgtatc ttataaaatc taatgcaat aattttcaga 1140
 ctggaatggg gtcgtggat tcagaggatc attgtttctt agttgcattt gatgtttgat 1200
 gagacttttc tcatgtgaaa cgtttattaa aacttcaaaa tggtgcaat gttgtcaaaa 1260
 aaaaaaaaaa 1268

<210> 2

<211> 240

<212> PRT

<213> Zea mays

<400> 2 Met Gly Arg Gly Lys Ile Val Ile Arg Arg Ile Asp Asn Ser Thr Ser

1 5 10 15

Arg Gln Val Thr Phe Ser Lys Arg Arg Asn Gly Ile Phe Lys Ala
 20 25 30

Lys Glu Leu Ala Ile Leu Cys Asp Ala Glu Val Gly Leu Val Ile Phe
 35 40 45

Ser Ser Thr Gly Arg Leu Tyr Glu Ser Ser Thr Ser Met Lys Ser
 50 55 60

Val Ile Asp Arg Tyr Gly Lys Ala Lys Glu Glu Gln Gln Val Val Ala
 65 70 75 80
 Asn Pro Asn Ser Glu Leu Lys Phe Trp Gln Arg Glu Ala Ala Ser Leu
 85 90 95
 Arg Gln Gln Leu His Asn Leu Gln Glu Asn Tyr Arg Gln Leu Thr Gly
 100 105 110
 Asp Asp Leu Ser Gly Leu Asn Val Lys Glu Leu Gln Ser Leu Glu Asn
 115 120 125
 Gln Leu Glu Thr Ser Leu Arg Gly Val Arg Ala Lys Lys Asp His Leu
 130 135 140
 Leu Ile Asp Glu Ile His Asp Leu Asn Arg Lys Ala Ser Leu Phe His
 145 150 155 160
 Gln Glu Asn Thr Asp Leu Tyr Asn Lys Ile Asn Leu Ile Arg Gln Glu
 165 170 175
 Asn Asp Glu Leu His Lys Lys Ile Tyr Glu Thr Glu Gly Pro Ser Gly
 180 185 190
 Val Asn Arg Glu Ser Pro Thr Pro Phe Asn Phe Ala Val Val Glu Thr
 195 200 205
 Arg Asp Val Pro Val Gln Leu Glu Leu Ser Thr Leu Pro Gln Gln Asn
 210 215 220
 Asn Ile Glu Pro Ser Thr Ala Pro Lys Leu Gly Leu Gln Leu Ile Pro
 225 230 235 240
 250

<210> 3
 <211> 1502
 <212> DNA
 <213> Zea mays

<400> 3
 ctgaagctgt cgtgttcctc acgtgcattg gctacaaaatg tatgtaacat agtaaatctg 60
 attgaaaaatg taaaatcttaa taattattat tacgacaaaac taaaaaccta ttccattttct 120
 ttgcataatga aaaatacagg tattccatgt aatataggaa caaatgtata atgcggctac 180
 tttgatgaac aatttataaa tgcataaaaa ggtatgtaac ataataaaatc cgtcatgtaa 240
 tggatgtact atataaaata gcaacataaa tttatgtca actaaaaacc gcctttttct 300
 atgcgaaaata atttacacta gtttcatata ggaatagatg gaaaacccaaa tcattatgt 360
 ctgcattattt atgcattattt ctggatgtcaa gaaagagcgc gtatagccag catgacagat 420
 ccattgttta caacaaatgt tcaaagggtgg taaagtaaag acctcgaccg gaaccattgg 480
 gcccactgtc aaaatttaat ggtgattcta catctaaaat gtttttttc aaatataaga 540
 caaatggttt gtatctttgc ttttacccca atgggttcta tcaatgtatgg ttgaggccct 600
 ccattatcta ggtatgtggc aagtgcattca acgaatatga tcacttctac actcgacgg 660
 tctccttaat ttatcacatgtt ttacagttat atatttatag ttatataatgg atgataacttt 720
 aaatggaaatc agaaatagaa tgcaatgtttt aaataaaacc aatgaatctt aaagaaaatgat 780
 ttggatccag atatagtccaa gcaactttagt acaatgtatg acatataatg ccaatgccaa 840
 catgtttcaa gtgttttaaa aaactatgtaa cattttctgc ttctgttactt caaatacgtaa 900
 aaatagcact gttttataat aagtacatctt attatcacctt ataaaaactttt tttccagcaac 960
 attttatataa ttgtgcatttca attaattttta acaacactta taaatttagtca caaatacgtaa 1020

acaaagcgcg cccatggcg cgctcaagct actagttttt atacagtgal aacaatacta 1080
ctatattaga cataaaattt ggtgtatgtt atttattttac tggtactact ttttaagtc 1140
ataatttagt gtactagaag ctactacgccc ctgacggaga caagcacgca caatagaaaa 1200
gagcgccctc gtgtccagag tgctcgccctt tagtttcgtg gctactatat actagctg 1260
acgattttta ttttttttctt ctcttgggtt atcattagca cggccattta gggtaatga 1320
gcagcgccca ataaacaaaag cccttgggtt gccaatcggg tgtcggcg cgcgcgcgc 1380
cgcgcttgc ctccgtccctt cctccctccctt cccttgcta aataaaagccaa gtagcagata 1440
ggggaaagtc tgctcgcaag ttgcateccgc cctgcgccaa gaaaagccat cgttcttccc 1500
ac

<210> 4
<211> 502
<212> DNA
<213> Zea mays

<400> 4	ctgaagctgt cgtgtcctc acgtgcattg gctacaaaatg tatgtaacat agtaaatctg 60
attgaaaagt taaatcttaa taattattat tacgacaaaac taaaaaccta ttccatttct 120	
ttgcataatgtt aaaaatacagg tattccatgt aatataggaa caaatgtata atgcggctac 180	
tttgatgaaac aatttataaa tgccattaaaa ggtatgtaaac ataataaaatc cgtcatgtaa 240	
tgttgtaact atataaaaata gcaacataaaa tttaattgca actaaaaacc gccttttct 300	
atgcgaaaata atttacacta gcttcatata ggaatagatg gaaaacccaaa tcattaatgt 360	
ctgcattatc atgcattattt ctggatgtgaa gaaagagcgc gtatagccag catgacagat 420	
ccattgttta caacaaaatgt tcaaagggtgg taaagtaaag acctcgaccg gaaccattgg 480	
gccccacttgc aaaattttaat gg 502	

<210> 5
<211> 500
<212> DNA
<213> Zea mays

```

<400> 5
tgatttacata ctaaaaatgt ttgtttcaa atataagaca atatggtgtt atctttgtt 60
ttaccccaat gggttctatc aatgatgggtt gaggtcctcc attattttagg atgtggtaa 120
gtgcataac gaatatgatc acttctacac tcgacgggtt tcctaaattt atacagtatt 180
acagttatata ttatatagtt atatattttat gatactttaa atggaaatcag aaatagaatg 240
caatgcttaa ataaaaaccaa tgaatcttaa agaaaagattt ggatccagat atagtccagc 300
aactttagac aatagtatac atataatgcc aatgccaaca tgtttcaagt gctcttaaaa 360
actatgtaca ttttctgtt cgtttactca aatacgtaaa atagcactgc ttctataataa 420
gtacatctat tatcacctat aaaactttt ccagcaacat ttatataatt tgcgtatcat 480
taattttaaac aacacttata

```

<210> 6
<211> 1000
<212> DNA
<213> Zea mays

actacgcctt gacggayaca agcacgcaca atagaaaaga ggcgcctcg 720
 ctcgccttta gtttctggc tactatatac tagctgcac gatttttatt ttctttctct 780
 tctggctat cattagcagc gccatttagg ggtaatgagc agcggccat aaacaaagcc 840
 ctggggtagc caatcggtt tcgcccgtcg cgccgcgcg cgtctggct cctgccttcc 900
 tcctccctt ctttgctaaa taaagcaagt agcagatagg ggaaagtctg ctgcgaagtt 960
 gcatccgccc tgcgccaaga aaagccatcg ttcttcccac 1000

<210> 7
 <211> 740
 <212> DNA
 <213> Zea mays

<400> 7
 tgaatcttaa agaaaagatt ggatccagat atagtcac aacttatgac aatagtatac 60
 atataatgcc aatgccaaca tgtttcaagt gctcttaaaa actatgtaca ttttctgtct 120
 cgtttactca aatacgtaaa atagcactgc ttatataaa gtacatctat ttcacccat 180
 aaaactttttt ccagcaacat ttatataatt tgcgtatcat taattttaac aacacttata 240
 aattagctca aatacgtaac aaagcgcgc catggggcg 300
 acagtgaaaa caataactact atattagaca taaaattagg tgcgtatcat ttatataactg 360
 gtactacttc tttaagtcat aatttaggtgt actagaagct actacgcctt gacggagaca 420
 agcacgcaca atagaaaaga ggcgcctcg gtcagagtg ctgcgccttta gtttctggc 480
 tactatatac tagctgcac gatttttatt ttctttctct tctggctat cattagcagc 540
 gccatttagg ggtaatgagc agcggccat aaacaaagcc ctggggtagc caatcggtt 600
 tcgcccgtcg cgccgcgcg cgtctggct cctgccttcc tctccctt ctttgctaaa 660
 taaagcaagt agcagatagg ggaaagtctg ctgcgaagtt gcatccgccc tgcgccaaga 720
 aaagccatcg ttcttcccac 740

<210> 8
 <211> 500
 <212> DNA
 <213> Zea mays

<400> 8
 aattagctca aatacgtaac aaagcgcgc catggggcg 300
 acagtgaaaa caataactact atattagaca taaaattagg tgcgtatcat ttatataactg 120
 gtactacttc tttaagtcat aatttaggtgt actagaagct actacgcctt gacggagaca 180
 agcacgcaca atagaaaaga ggcgcctcg gtcagagtg ctgcgccttta gtttctggc 240
 tactatatac tagctgcac gatttttatt ttctttctct tctggctat cattagcagc 300
 gccatttagg ggtaatgagc agcggccat aaacaaagcc ctggggtagc caatcggtt 360
 tcgcccgtcg cgccgcgcg cgtctggct cctgccttcc tctccctt ctttgctaaa 420
 taaagcaagt agcagatagg ggaaagtctg ctgcgaagtt gcatccgccc tgcgccaaga 480
 aaagccatcg ttcttcccac 500

<210> 9
 <211> 410
 <212> DNA
 <213> Zea mays

<400> 9
 taaaattagg tgcgtatcat ttatataactg gtactacttc tttaagtcat aatttaggtgt 60
 actagaagct actacgcctt gacggagaca agcacgcaca atagaaaaga ggcgcctcg 120
 gtcagagtg ctgcgccttta gtttctggc tactatatac tagctgcac gatttttatt 180
 ttctttctct tctggctat cattagcagc gccatttagg ggtaatgagc agcggccat 240
 aaacaaagcc ctggggtagc caatcggtt tcgcccgtcg cgccgcgcg cgtctggct 300
 cctgccttcc tctccctt ctttgctaaa taaagcaagt agcagatagg ggaaagtctg 360
 ctgcgaagtt gcatccgccc tgcgccaaga aaagccatcg ttcttcccac 410

<210> 10

<211> 310
 <212> DNA
 <213> Zea mays

<400> 10
 aaacgcacac atagaagcat cattccccctc tcggctagct ccctccctctc tccctccctct 60
 tcgtcttccct ctttcccttccct cccttgggaa acctgctgcc tttgagcttt ctttctcgag 120
 tgctccacc agatctccctc ctcccttaccc ttttggcac gttcggcggc gcgcgcgag 180
 atagatagat cccgcacatcg tcgtcgccgt ctttgcctcc gatcgaggg ccacaaccac 240
 aacctctcgc tcccttagcgt gcgcgcgca gccagggtca agaagagagc tagctatagg 300
 ccggagatcg 310

<210> 11
 <211> 251
 <212> DNA
 <213> Zea mays

<400> 11
 agaagagataa aactgccgtc ttatgatgct gaaggaaact atttattgtg aagagatgat 60
 actcagagaa agacatattt gtggcaggga gatttggat atgaacttat aaatgtaatg 120
 caaataatattt tcagactgga atggggtcgt ggaattcaga ggatgattgc tttctagttg 180
 catttgcgtt ttatgagac ttttctcatg tgaaacgttt attaaaactt caaaatggtg 240
 cagatgttgc t 251

<210> 12
 <211> 5031
 <212> DNA
 <213> Zea mays

<220>
 <221> misc_feature
 <222> (1813)..(1994)
 <223> EXON I

<220>
 <221> misc_feature
 <222> (2871)..(2954)
 <223> EXON II

<220>
 <221> misc_feature
 <222> (3202)..(3258)
 <223> EXON III

<220>
 <221> misc_feature
 <222> (4066)..(4166)
 <223> EXON IV

<220>
 <221> misc_feature
 <222> (4248)..(4289)
 <223> EXON V

<220>
 <221> misc_feature
 <222> (4367)..(4449)
 <223> EXON VI

<220>

<221> misc_feature
<222> (4523)..(4676)
<223> EXON VII

<220>

<210> 13
<211> 20
<212> DNA
<213> Zea mays

<400> 13
atggggaggg gaaggattga

20

<210> 14
<211> 21
<212> DNA
<213> Zea mays

<400> 14
ctgttgtgg cgtactcgta g

21

<210> 15
<211> 21
<212> DNA
<213> Zea mays

<400> 15
aaaaaacaada tcqacatcaa g

21

<210> 16
<211> 19
<212> DNA
<213> Zea mays

<400> 16
gkcgwactcg tagaggcg

19

<210> 17
<211> 20
<212> DNA
<213> Zea mays

<400> 17
atggggaggg gacgmggtga

20

<210> 18
<211> 20
<212> DNA
<213> Zea mays

<400> 18
gctgccgaac tcgttagagct

20

<210> 19
<211> 21
<212> DNA
<213> Zea mays

<400> 19
gttgttggcg tactcgtaga g

21

<210> 20
<211> 20
<212> DNA
<213> Zea mays

<400> 20
gggcgcaagg tacagctgaa

20

<210> 21
<211> 21
<212> DNA
<213> Zea mays

<400> 21
gttgttggcg tactcgtaga g

21

<210> 22
<211> 21
<212> DNA
<213> Zea mays

<400> 22
aagggcaaga ctgagatcaa g

21

<210> 23
<211> 23
<212> DNA

<213> Zea mays

<400> 23
cacttgaact cttttacgct tat

23

<210> 24
<211> 23
<212> DNA
<213> Zea mays

<400> 24
gacaatcttg acacatgtat gaa

23

<210> 25
<211> 22
<212> DNA
<213> Zea mays

<400> 25
cgattcaaat cgtgaatctc at

22

<210> 26
<211> 20
<212> DNA
<213> Zea mays

<400> 26
cccccgggct gcaggaattc

20

<210> 27
<211> 22
<212> DNA
<213> Zea mays

<400> 27
tttttagcaac atctgcacca tt

22

<210> 28
<211> 19
<212> DNA
<213> Zea mays

<400> 28
tctcggctag cttectct

19

<210> 29
<211> 30
<212> DNA
<213> Zea mays

<400> 29
cggcctatacg cttagctctt tcttgaccct

30

<210> 30
<211> 27

<212> DNA
<213> Zea mays

<400> 30
agggtcaaga agagagctag ctataagg

27

<210> 31
<211> 28
<212> DNA
<213> Zea mays

<400> 31
gctaaggagc gagaggttgtt ggtttgtgg

28

<210> 32
<211> 26
<212> DNA
<213> Zea mays

<400> 32
ccacaaccac aacctctcgc tcctta

26

<210> 33
<211> 19
<212> DNA
<213> Zea mays

<400> 33
tcggcagttt acgggagat

19

<210> 34
<211> 20
<212> DNA
<213> Zea mays

<400> 34
tttagcaactc attatacgac

20

<210> 35
<211> 20
<212> DNA
<213> Zea mays

<400> 35
agaaaaccaga gatgttccag

20

<210> 36
<211> 21
<212> DNA
<213> Zea mays

<400> 36
caacatctgc accattttga a

21

<210> 37

<211> 23
<212> DNA
<213> Zea mays

<400> 37
gtcacatcaa caagcgtcct acc

23

<210> 38
<211> 23
<212> DNA
<213> Zea mays

<400> 38
taagagcatc aacagtagta acg

23

<210> 39
<211> 20
<212> DNA
<213> Zea mays

<400> 39
agaaaccaga gatgttccag

20

<210> 40
<211> 21
<212> DNA
<213> Zea mays

<400> 40
caacatctgc accattttga a

21